

Reputation in Self-Organized Communication Systems and Beyond

(Invited Paper)

Jochen Mundinger
EPFL-IC-LCA
BC203, Station 14
CH-1015 Lausanne

jochen.mundinger@epfl.ch

Jean-Yves Le Boudec
EPFL-IC-LCA
BC203, Station 14
CH-1015 Lausanne

jean-yves.leboudec@epfl.ch

ABSTRACT

Efficiently handling reputation is important in dealing with free-riding, malicious attacks and random failures in self-organized communication systems. At the same time, work in this context is often found to be relevant in many other disciplines, in particular the social sciences. A number of distributed reputation systems have been proposed and analyzed, although research has not been very coherent. In this paper, for the first time, we provide an overview of the state-of-the-art in the various computer science communities as well as the social sciences. In particular, we present results obtained from our mathematical model devised to investigate the impact of liars on their peers' reputation about a subject. We find that liars have no impact unless their number exceeds a certain threshold (phase transition). We give precise formulae and quantify the impact, thereby providing insight into fundamental questions in social networks as well as facilitating performance evaluation and optimization of distributed reputation systems in communication networks. We conclude by suggesting fundamental directions for future research into reputation.

Categories and Subject Descriptors

C.2.1 [Computer-communication networks]: Network architecture and design—*Distributed networks*; D.2.4 [Computer-communication networks]: Distributed systems—*Distributed applications*; I.2.11 [Distributed Artificial Intelligence]: Multiagent systems, Intelligent agents; J.4 [Social and behavioral sciences]: Sociology, Economics; C.4 [Performance of systems]: Modelling techniques, performance attributes; G.3 [Probability and Statistics]: Markov processes, stochastic processes

General Terms

Algorithms, Performance, Theory, Economics, Security

Keywords

Reputation, Trust, Liars, Bayesian belief propagation, Social Networks, Interacting particle system, Statistical physics model, Discrete event dynamic system, Phase transition, Performance optimization

1. INTRODUCTION

Self-organized computer and communication systems have received increasing attention over the last few years, in terms of both deployment and research. They are typically organized according to the peer-to-peer (P2P) organization principle. That is, participants in the system are equals in that they have equivalent capabilities and responsibilities – they are peers. Such P2P systems can also be found in a variety of other networks, such as social or biological networks. Thus it is not surprising that there is a wealth of problems that is also of interest in other disciplines.

One of the major problems in self-organized communication systems is that of cooperation. Typically, users are concerned primarily about their own benefits and thus cooperation and fairness cannot be guaranteed. This selfish behaviour is called *free-riding* and is a well-known phenomenon in economics. Malicious attacks and random failures are other important problems. A promising approach to all of these is that of using reputation systems.

Reputation itself has been considered in many disciplines other than computer and communication science, including economics, sociology, psychology, management science as well as marketing. For a number of broad-ranging studies on reputation the reader is referred to [35].

In computer science, the idea is comparatively new. However, with the increasing popularity of self-organized communication systems, distributed reputation systems in particular have received increasing attention in the last few years. Even within computer science research activities have not been very coherent, though, and have almost evolved separately in the artificial intelligence, Internet-based P2P and Mobile Ad-Hoc Networks communities. In fact, there is not even a consistent definition of *reputation* itself, and, closely linked, that of *trust*.

As a convention, following the Oxford English Dictionary, we shall adopt that reputation is an estimate about a person's actual quality. *Person* is the appropriate term for social networks. In the context of computer networks we shall replace it with *user* (of the system), *node* (in the network) or simply a *peer*. Similarly, *quality* refers to the behaviour that is of interest in a given context.

It is beyond the scope of this paper to provide an overview of reputation research in all the disciplines mentioned above. In addition to computer science, we shall thus focus on the social sciences, perhaps the most generic discipline for the study of reputation. For the first time, we provide an overview of the state-of-the-art in the different computer science communities concerned with reputation as well as the social sciences.

In particular, we will present results obtained from our mathematical model devised to investigate the impact of liars on their peers' reputation about a subject. The model can be viewed as a generalization of the voter model [37], a well-known interacting particle system in statistical physics. It is appropriate both for communication networks and social networks, although the results are viewed from a slightly different point of view. We find that liars have no impact unless their number exceeds a certain threshold, that is we observe a phase transition behaviour. We give precise formulae and quantify the impact, thereby providing insight into social networks as well as facilitating performance evaluation and optimization of distributed reputation systems in communication networks.

Historically oriented, the overview is organized according to the almost separate evolution in the different communities concerned with reputation. However, we put our model into the context of related work in all the different communities and argue for a more coherent terminology and approach. We also argue that, while implementation details are crucial, mathematical models should be considered as a means to help answer questions of a more fundamental nature.

The rest of this paper is organized as follows. We provide background for self-organized communication systems and motivate the use of reputation systems in Section 2. The overview of the state-of-the-art in computer science can be found in Section 3 and our mathematical model for liars' impact on reputation is discussed in Section 4. We then turn to related work on reputation in the social sciences (Section 5) and conclude by suggesting fundamental directions for future research into reputation in Section 6.

2. SELF-ORGANIZED COMMUNICATION SYSTEMS

There are many examples of P2P communication systems, such as Internet-based P2P systems and Mobile Ad-Hoc Networks. Other examples include Weblogs (or *blogs*), *pod-casting* and Wikipedia¹, although there is an element of centralization, whereas we shall be concerned with fully distributed systems. We now look at Internet-based P2P systems and Mobile Ad-Hoc Networks in turn.

¹<http://wikipedia.org/>

2.1 Internet-Based P2P Systems

In this context, P2P refers to systems in which the users function simultaneously as both clients and servers to their peers in the network. This differs from client/server (C/S) architectures, in which some users are dedicated to serving the others. Each user is running software that implements the same communication protocols and resource sharing techniques. The current Internet being essentially client/server, systems achieve the P2P organization in the form of application-level overlay networks. A basic feature is that users share local resources with their peers. Often, these are computer resources such as processing power, cache or disk storage. However, services or information in form of files are also shared.

In fact, the Internet itself was originally conceived as a P2P network in that even though individual applications were of C/S type, the user behaviour as a whole was entirely symmetric. Recently, however, P2P has been rediscovered as an attractive paradigm for building distributed networked applications. The popularity of file sharing systems like Napster, Kazaa and more recently the BitTorrent protocol² demonstrate the development (back) towards the symmetric architecture of equal peers. In terms of usage and popularity, [33] find that a significant proportion of Internet traffic arises from P2P systems and that this is likely to increase further in the future.

2.2 Mobile Ad-Hoc Networks

Mobile Ad-Hoc Networks are composed of equal participants, the nodes, which communicate in a decentralized fashion over wireless channels. Typically multi-hop rather than direct communication between nodes is considered. For example, the Terminode Project³ envisages a wireless network of small personal devices owned by everyone in a wide area [30].

Mobile Ad-Hoc Network are also organized according to the P2P principle. They are autonomous (independent of any infrastructure), self-organized and decentralized. Moreover, there are additional issues of mobility, wireless links, limited battery power and the important resource shared is the forwarding of packets.

2.3 Problems and Solution Approaches

While such self-organized systems have many important advantages such as scalability to potentially large number of users, there are also problems.

In particular, in most applications users are individuals that are primarily interested in their own benefit. As there is a natural incentive for users to only consume, but not contribute, cooperation and fairness cannot be guaranteed. In Internet-based P2P networks, users might not want to provide bandwidth. In Mobile Ad-Hoc Networks, the users might not want to provide their own limited battery to forward other users' packets. The ability to receive messages is often not sufficient motivation to actively provide services. Note that technical competence of the users is not a necessity. If a company were to produce devices that did not

²<http://bitconjurer.org/BitTorrent/protocol.html>

³<http://www.terminodes.org>

cooperate and therefore had a longer runtime, that might be attractive to customers.

This behaviour is called *free-riding* and is a well-known phenomenon in economics. It often occurs in situations arising in the context of public goods [18]. A public good is both *non-rivalrous*, meaning that consumption by one does not limit consumption by others, and *non-exclusive*, meaning peers cannot be excluded from the benefits. In our case, all users benefit from the service, however, without necessarily contributing themselves. They free-ride in that they increase their utility by taking more than their fair share of the benefits or rather, by not shouldering their fair share of the costs. The *free-rider problem* is that as a result this service might not be provided at all or without sufficient quality of service [53].

The free-rider problem often occurs in everyday life. For example, consider air pollution, logging of forests, over-fishing of the oceans or private vehicles jamming public roads. A famous illustration of the free-rider problem is the *Tragedy of the Commons* [27].

Effects can be detrimental. In Internet-based P2P networks, for example, they have been illustrated in [3], [55] and [17].

The loss in social welfare achieved by such uncoordinated individual utility maximizing behaviour compared to the social optimum is sometimes referred to as the *price of anarchy*. This term was originally introduced in the context of delays in Internet flows [48], higher delay being bad, but can easily be adapted to other metrics of interest. It is typically used in game theoretic models where it refers to the ratio between the worst-case Nash equilibrium welfare and the optimal social welfare [50].

Although altruistic behaviour has been observed, it is not clear to which extent this will help in self-organized communication systems. *Altruism* is the practice of being helpful to other people with little or no interest in being rewarded for one's efforts. The concept has a long history in philosophical and ethical thought. For recent work on altruism in Economics see [4] and [24]. Also see [47].

Thus, a number of mechanisms have been proposed, ranging from incentive mechanism to reputation systems and artificial immune systems. They need to take into account both the economic side and the engineering side. For example, identity is an issue in all these systems [25, 23]. We shall consider them in turn.

2.3.1 Incentive mechanisms

Incentive mechanisms are aimed at making it advantageous for users to act in such a way that the resulting social welfare is optimal [7].

The standard approach here is that of accounting schemes or pricing mechanisms which have been applied successfully in rate control in wireline networks, resource control in wireless networks as well as in the wider context of communication networks [19]. This might involve payments in kind or virtual or real payments. For example, see [16] for a virtual currency called nugglets and a nugglet counter. [21] suggest

a pricing scheme as used for Internet traffic based on notional credit, all in the context of Mobile Ad-Hoc Networks. [26] examine micropayment mechanisms in P2P file sharing networks.

Pricing often leads to rather complex mechanisms. Thus, rules are considered as an alternative [5]. That is, actions of users are constrained locally by the software. Rules might be more appropriate for simple symmetric P2P systems of low cost resources. A simple example of a rule is to force users to contribute in order to consume. Rules are used by [20] in the context of P2P Wireless LAN Consortia. Incentive mechanisms typically do not have an enforcement component.

2.3.2 Reputation systems

Two other problems often incurred in P2P networks are malicious attacks and random failures. Reputation systems address both these issues as well as incentive problems. Here, users keep track of their peers' behaviour and exchange this information with others in order to compute a reputation value about their peers. Users with a good reputation are then favoured.

Reputation systems have proven useful and are popular in online auctioning systems such as eBay [58] or online book stores such as Amazon and can be viewed as a substitute for the word-of-mouth mechanisms in social networks. Correspondingly, current research is concerned with investigating the use of fully distributed reputation systems in self-organized communication systems. However, the distributed nature leads to potentially very complex behaviour that needs to be understood better. We will provide an overview of the state-of-the-art in distributed reputation systems in Section 3.

2.3.3 Artificial immune systems

More recently, artificial immune systems have begun to be considered. They are aimed primarily at misbehaviour detection and designed so that they adapt to normal behaviour, but also recognize new misbehaviour patterns that had not been anticipated in the system design phase [57]. Moreover, artificial immune systems use mechanisms for faster detection of repeated misbehaviour [34, 54]. An important potential advantage of such systems is their inherent randomness that provides diversity at the population level. Even if some computers are vulnerable to an attack, there should be many others that are resistant to the same attack. However, implementations seem to depend very much on the particular application and at the moment there does not appear to be a working application in which an artificial immune system has proved superior to other approaches.

3. REPUTATION SYSTEMS IN COMPUTER SCIENCE

As seen in the previous section, using reputation systems is a promising approach to incentive problems as well as malicious attacks and random failures in self-organized computer and communication systems. Indeed, a number of reputation systems have been proposed and we will now provide an overview of the state-of-the-art.

Due to space restrictions, we shall focus on fully distributed reputation systems in this paper. For reputation systems relying on a centralized component, the reader is referred to [49, 22, 38]. A typical application scenario for those reputation systems are online trading/online auction mechanisms such as eBay.

In a fully distributed reputation system, users keep track of their peers' behaviour and exchange this information directly with others rather than with the help of some centralized entity. Each user merges their own first hand information with the second hand information they receive in order to compute a reputation value about each of their peers. This might be an automated procedure. Users with a good reputation are then favoured.

The advantage of a reputation system over merely using first hand information is two-fold. Firstly, an accurate estimate of some subject's behaviour can be obtained faster. Secondly, a user can have a reputation value about a subject without ever having interacted with it himself.

However, an inherent problem with any such mechanism is the vulnerability to liars. Some user might have an interest in spreading false information, so naively believing all second hand information is problematic. Reputation values must be accurate at least to some degree and thus robust against liars.

Reputation systems have been considered, almost separately, in different computer science communities. Although it is a relatively recent area, research efforts have increased significantly over the last few years. We will now consider artificial intelligence, mobile-ad hoc and P2P literature in turn.

It should be noted that we organize our survey in this manner, because *historically*, research has evolved this way, not because we think it particularly suitable. In fact it is not. A number of ideas have shown up in several communities, but explicit links between them are rare. Sometimes there are more links to other disciplines such as the social sciences than to other communities within computer science. Moreover, the terminology is rather inconsistent. Often, there are actually different concepts, but occasionally what some authors call *reputation*, others call *trust*. Introducing our model in Section 4, we link it with related work from all these communities and argue that this is a more efficient organization for reputation systems research (Section 6).

3.1 Artificial Intelligence Research

A review on reputation in computer science with a focus on the artificial intelligence literature is given in [52]. It is mainly concerned with implementations, as such does not consider theoretical models.

In their terminology, *reputation* is one of the elements that help to build *trust*. Reputation systems are classified according to (A1) conceptual model, (A2) information sources (direct experiences, witness information, sociological information (A3) prejudice (conclusions drawn from group membership), (A4) visibility types, (A5) granularity, (A6) agent behaviour assumption (honest, lie partially, lie and mechanisms), (A7) type of exchanged information (discrete, con-

tinuous) and (A8) reliability measure. It is observed that using sociological information in (A2) about the links between people (such as competition, collaboration) is not considered by many systems. [51] uses this by employing sociograms, that is, graphs representing relational data. Although, it is difficult to construct the sociograms – and this is not discussed in the paper – the approach is to linking different disciplines. Similarly, the prejudice information based on group memberships has not been considered much.

Another review with a focus on online service provision is given in [31]. To them, *reputation* is, somewhat loosely defined, what is generally said or believed about a person's character or standing. (*Reliability*) *trust* is defined by the subjective probability by which a person expects that a peer performs a given action. They consider both centralized and distributed reputation systems.

[62] and [59] also belong in this context. As they are related to our model in Section 4, we discuss them there. For other technical papers, the reader is referred to the references in the reviews above.

3.2 P2P Systems Research

A number of reputation mechanisms have been suggested and studied. A comprehensive survey and more detailed overview of reputation systems suggested for Internet-base P2P systems can be found in [2].

Here, *trust* is the extent to which a user trusts a peer behaves well (cooperates) and *reputation* is the commonly shared believe how likely a peer behaves well (cooperates). They distinguish approaches according to (B1) social networks (by which they mean graph theoretic models that consider transitivity of trust along the edges), (B2) probabilistic estimation, (B3) game theoretic models (both classical and evolutionary game theory).

[1] suggest a mechanism for P-Grid, a P2P system, that spreads negative information only. The reader is referred to [32] for the EigenTrust algorithm, a method to compute global trust values in the presence of pre-trusted peers. Another mechanism is PeerTrust as introduced by [60]. Others can be found in the references of the survey.

3.3 Mobile Ad-Hoc Networks Research

A number of reputation mechanisms have been suggested and studied. A comprehensive survey and more detailed overview of reputation systems suggested for Mobile Ad-Hoc Networks can be found in [13]. Reputation systems are classified according to (C1) representation of information and classification, (C2) use of second-hand information, (C3) trust and (C4) redemption and secondary response.

The CONFIDANT Protocol (Cooperation of Nodes, Fairness In Dynamic Ad-hoc NeTworks) was proposed in [11]. Reputation is based on direct observations based on a neighbourhood monitor as well as second hand information from other nodes and are updated according to a Bayesian estimation. In addition, there is a trust manager and a path manager that implements the reaction by avoiding and isolation misbehaving nodes. It is demonstrated that using second-hand information can significantly accelerate the de-

tection and subsequent isolation of malicious nodes. The robustness of the system against wrong accusations is also considered [12].

The Collaborative REputation mechanism (CORE) was introduced in [39] with a game theoretic analysis. Each node of the network monitors the behaviour of its neighbours with respect to a requested function and collects observations about the execution of that function. Reputation takes into account subjective observations, indirect reports from peers and functional reputation that is task-specific. Based on the collected observations, it computes a reputation value for each neighbour. Selfish nodes are also avoided and service denied.

OCEAN [9] and SORI [28] are also discussed in more details in the survey. These systems have all been developed for a fairly specific set of assumptions, in particular assuming Dynamic Source Routing (DSR) [29].

4. MODELLING THE LIARS IMPACT

As described in the previous section, there is a tradeoff between speed and accuracy: the more second hand information is used, the faster an estimate of some subject's behaviour can be obtained, however, the more vulnerable it is to liars. In order to be useful, reputation values need to be accurate, at least to some degree, though. In this section, we will present results obtained from our mathematical model devised to investigate the impact of liars on their peers' reputation about a subject.

Our starting point is a simple and fully distributed idea to address the problem of liars. We suppose, a user believes second hand information if and only if it does not differ too much from the user's reputation value.

This idea, called the deviation test, was used as one component in the reputation system [12] in the context of Mobile Ad-Hoc Networks (cf. Section 3.3). The other, much more complex component allows the use of second hand information from trusted peers. To this end, each user maintains both a reputation and a trust value about each of his peers. As opposed to reputation, trust values are based on compatibility and thus indicate agreement. However, in simulations the deviation test on its own was found to perform surprisingly well. It seemed intriguing that such a simple idea works so well and this motivated us to analyze it in more detail and in a wider context.

We consider an abstract model of a reputation system in order to address fundamental questions independently of the details of implementation. As such, we are not concerned with the detection and response components of a system, but focus on the formation of reputation. The detection component depends on the application scenario and we assume that misbehaviour can be told apart from good behaviour. Moreover, we assume that if reputation values can be computed accurately, then there exists a response mechanism using them to obtain the desired effects. Typically, this might mean exclusion of the misbehaving user from benefits.

In general, robustness against liars has not been analyzed in depth, although some related work can be found in the

artificial intelligence community (cf. Section 3.1). [62] also consider the problem of liars via some models of deception. Their approach is based on the weighted majority technique where the last second hand information is tested by comparing it to next direct interaction. The analysis is based only on simulation. [59] is also concerned with filtering out manipulated second hand information that seems unlikely. However, they consider quantiles of the Beta distribution rather than distance. This paper, too, is merely based on simulation. In the context of centralized reputation systems, [40] consider incentive mechanisms not to stimulate good behaviour in the network (cf. Section 2.3.1), but to stimulate honest reports within the reputation system. They provide a game theoretic analysis to show that honest reporting is a Nash equilibrium.

The abstract nature of our model also makes it relevant in a much wider context. One of them is social networks and we will explain this as well as related work in Section 5.

One of the fundamental questions we address about the comparison of two different scenarios as regards to the second hand information. In the first one, *Reputation* – based on all previous observations including indirect ones – is passed on as second hand information. In the second one, only *Direct Observations* are passed on as second hand information.

In this paper, we only provide an overview. The reader is referred to [45] and [44] for the analysis of the model for two users, the former considering a one-dimensional simplification. Together with [43], these cover self-organized communication networks. For social networks and the generalization to a network of N people see [46]. Full details and derivations can be found in [42].

We outline our methodology and summarize the modelling assumptions in Section 4.1. In Section 4.2 we illustrate the model using a typical simulation sample path. The main results are summarized in Section 4.3.

4.1 Methodology

Our model is a stochastic process formulation based on a number of assumptions. We suppose that there is a single subject under consideration, behaving positively with and negatively with probabilities θ and $1 - \theta$ respectively, and independently. Note that in our model reputation values of several subjects do not interact, so the general case can be decomposed into multiple instances of our model.

We suppose that there are N_h honest users and N_l liars in the network and each user i has counters (x_n^i, y_n^i) for positive and negative information respectively. The reputation values are obtained as $R_n^i = x_n^i / (x_n^i + y_n^i)$ in $[0, 1]$, 1 being positive and 0 being negative. Counters and corresponding reputation values are updated with events in a Bayesian fashion. However, whereas direct observations are always accepted, second hand information is accepted by a user only if considered likely, i.e. only if it does not differ by more than a threshold Δ from the user's current reputation value. Even if accepted, it is weighted by a factor ω_{weight} .

Moreover, we include discounting with a factor ρ , so that old observations gradually become less important. Liars

are assumed to naively report extremely negative reputation values and observations respectively. The positive case is similar by symmetry.

Interactions are assumed to happen symmetrically according to a Poisson process framework. Thus, a given interaction is a direct observation with probability p , indirect from a liar with probability q and from an honest peer with probability $r = 1 - p - q$ where these probabilities depend on the number of liars in the network.

The stochastic process formulation can be viewed as a generalization of the well known voter model [37]. With $\rho = 0$, $\omega = \Delta = 1$, $p = q = 0$, $r = 1$ and initial value (x_0^i, y_0^i) either $(1, 0)$ or $(0, 1)$ for all i , we recover the voter model on a complete graph.

From the stochastic process formulation, we derive an ordinary differential equation by averaging the dynamics and passing to a *fast-time scaling* limit. That is, we scale time so that events occur more frequently, i.e. users make observations at a higher rate, but at the same time the impact of each observation is reduced by the same factor. We then derive the solutions of the differential equation and study their fixed points. Thus, our approach can be called a *mean-field* approach [61]. Moreover, we use simulation and direct computation to confirm the analytical results.

4.2 Typical Sample Path

For better illustration, we show a typical sample path obtained from the simulations with parameter set 1 in Figure 4.2. It is obtained in the scenario where *Direct Observations* only are passed on as second hand information.

PARAMETER SET 1. $\theta = 0.8$, $p = 0.2$, $q = 0.6$, $r = 0.2$, $\omega = 0.75$, $N_h = 100$, i.e. $N_l = (N_h - 1)q/r = 297$ liars, $\rho = 0.995$ and initial values $R_0 = 0$. $\Delta = 0.32$. We carry out $80000N_h$ steps. Thus, the fixed points obtained from analysis are $R_{true}^* = 0.8$ and $R_{false}^* = 0.2$, but, here, not R_{inter}^* (*Direct Observations*).

The upper and lower boundaries in the plot correspond to reputation values 1 and 0. The intermediate lines correspond to fixed point reputation values R_{true}^* , R_{inter}^* and R_{false}^* . Two kinds of average reputation values are plotted in black and grey, two individual reputation values of two users are plotted in blue and yellow. The individual reputation values increase and settle down at R_{false}^* before increasing further, one by one, past R_{inter}^* and settling down at R_{true}^* . This confirms existence as well as the values of the fixed points R_{false}^* and R_{true}^* .

4.3 Main Results

A main result is that, in order to have an impact, the number of liars N_l in the network needs to exceed a certain threshold. That is, there is a phase transition behaviour. Alternatively, this can be phrased in terms of the parameter Δ rather than in terms of N_l . If Δ is below a certain threshold, the liars have no impact.

We provide precise formulae for these critical values and quantify the impact. We find that *Reputation* and *Direct*

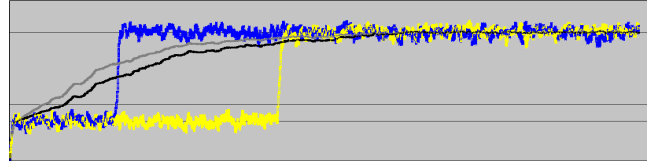


Figure 1: Typical graph of reputation plotted vs. time obtained from a simulation of parameter set 1. The upper and lower boundaries in the plot correspond to reputation values 1 and 0. The intermediate lines correspond to fixed point reputation values R_{true}^* , R_{inter}^* and R_{false}^* . Two kinds of average reputation values are plotted in black and grey, two individual reputation values of two users are plotted in blue and yellow. The individual reputation values increase and settle down at R_{false}^* before increasing further, one by one, past R_{inter}^* and settling down at R_{true}^* . This confirms existence as well as the values of the fixed points R_{false}^* and R_{true}^* (*Direct Observations*).

Observations coincide if and only if $\theta > 2\Delta$. For *Reputation*, second hand information does not improve accuracy, whereas for *Direct Observations* it does. We quantify this difference.

In the context of communication networks we can use our results to give guidelines for a good choice of parameters and hence system design. For example, for maximal gains in terms of speed and without compromising on accuracy, the system parameter Δ should be chosen as the threshold.

More specifically, a typical result is of the following form (*Reputation*).

THEOREM 1. Consider the scenario where Reputation is passed on as second hand information.

If $\Delta < \Delta_{c1} = (p\theta)/(p + q\omega)$,

$$(x, y) = \frac{1}{(1 - r\omega)(1 - \rho)} (p\theta, p(1 - \theta)) \quad (1)$$

is the unique fixed point of the mean ODE. It is asymptotically stable and all trajectories are attracted to it. The corresponding reputation value is

$$R_{true}^* = \theta. \quad (2)$$

If $\Delta_{c1} \leq \Delta < \Delta_{c4} = \theta$ there is a second, false fixed point

$$(x, y) = \frac{1}{(1 - r\omega)(1 - \rho)} (p\theta, p(1 - \theta) + q\omega). \quad (3)$$

Both are asymptotically stable, attracting trajectories from $x(t)/(x(t) + y(t)) > \Delta$ and $x(t)/(x(t) + y(t)) \leq \Delta$ respectively. The corresponding reputation value is

$$R_{false}^* = \theta p / (p + \omega q). \quad (4)$$

If $\Delta_{c4} \leq \Delta$, then only the latter, false one is asymptotically stable and all trajectories are attracted to it.

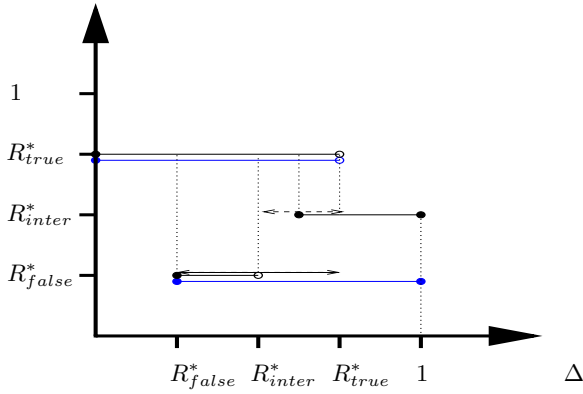


Figure 2: Bifurcation plot in terms of Δ : As Δ increases from 0 to 1 the number of fixed points changes. The case with *Reputation* is shown in blue, the case with *Direct Observations* is shown in black. *Reputation*: depending on the parameters, up to 2 fixed points, 2 critical values. *Direct Observations*: depending on the parameters, up to 3 possible fixed points, 4 critical values.

For *Direct Observations*, the corresponding result is slightly more complicated. We compare the two in Figure 4.3. As Δ increases from 0 to 1 the number of fixed points changes. The case with *Reputation* is shown in blue, the case with *Direct Observations* is shown in black. For *Reputation*, depending on the parameters, there are up to 2 fixed points and 2 critical values. For *Direct Observations*, depending on the parameters, there are up to 3 possible fixed points and 4 critical values.

So far, we have assumed independent subject behaviour. It will be interesting to consider the case when direct observations are correlated.

Another extension is to consider strategic lying, that is liars attempting something more subtle than simply telling extreme lies. For example, they could lie in some proportion of reports only or they could always report intermediate behaviour in an attempt to conceal their lies.

It will also be interesting to extend our results from the symmetric situation we have considered thus far to an asymmetric situation. In many social networks, people are not symmetric in terms of their interactions. Lai and Wong [36], for example, have examined the tie effect on information dissemination in the context of rumour spreading. They find that information transmitted via kin ties tends to arrive at the respondent faster than via non-kin ties or other communication channels. We turn to social networks in the next section.

5. REPUTATION RESEARCH IN THE SOCIAL SCIENCES

Social network analysis [56] is concerned with relationships of individuals in a society. The concept of reputation in social networks is a natural one and we experience it in everyday life.

Let us consider a dense, closed social network. By this we mean that everyone in it is connected to everyone else by similarly strong relationships. People in the network are assumed to take an interest in the behaviour of some subject which can be either positive or negative. They interact with this subject directly. They also interact with each other, e.g. in conversations, and thereby pass on their own experiences with the subject to their peers. Based on both direct experience and indirect information they form their reputation about the subject. An example is the social network of truck drivers interested in the quality of food of a highway restaurant. Alternatively, the subject might be part of the social network itself and there might be more than one subject. This is the case when people in the network gossip about each other.

Sharing experiences with one's peers serves the purpose of using information more efficiently: By also considering other people's experiences, one is able to get a more accurate idea about the actual subject behaviour faster. However, it might be the case that not every person in the social network passes on their experiences with the subject truthfully. There might be liars with an interest in manipulating their peers' reputation about the subject. In the absence of trust, these might distort the overall reputation of the subject in the network.

So the question arises whether second hand information is or should always be believed. We assume that this is not so. Rather, if a person is confronted with information that is not verifiable, they will probably believe it only if, to them, it seems likely. However, they will ignore it if, to them, it seems unlikely. Moreover, they will not necessarily attach the same weight to an experience reported by a peer compared to their own direct experiences. We also assume that people gradually forget experiences they have made a long time ago and that in the current reputation about the subject recent experiences are given greater impact as a result.

Then we recover our stochastic model from the previous section. There is the subject under consideration, behaving positively with and negatively with probabilities θ and $1 - \theta$ respectively, and independently. There are N_h honest people and N_l liars in the network and their opinion is represented by reputation values R_n^i in $[0, 1]$, changing with new experiences. The threshold parameter distinguishing between likely and unlikely reports is Δ and the weighting parameter is ω_{weight} , forgetting is accounted for via the discounting factor ρ . Again, the Poisson process framework with parameters p , q , and $r = 1 - p - q$ is a natural one to start with.

As in the communication networks context, the results tell us about the phase transition in the social networks context. Unless the number of liars exceeds a threshold, they do not have any impact on their peers reputation about the subject. Moreover, we find how much better it would be, if everyone only passed on their own direct experience rather than gossiping. In the context of social networks it is not that we are interested in optimizing a system parameter. Rather, we predict reputation given certain assumptions about people's behaviour, thereby offering social scientist a different approach to study reputation in social networks.

Similar questions have also been studied in the social sciences. [41], for example, examine to what extent does deteriorate a person's labour market position.

[15] investigates how the competitive advantage known as social capital depends on the structure of the social network. He also focuses on closed networks rather than networks of interdependent groups (brokerage, cf. also [14]) and evaluates two competing hypotheses. Firstly, the so-called bandwidth hypothesis that network closure enhances information flow. This is found in closure models of social capital and as well as in reputation models in economics. Secondly, the so-called echo hypothesis that closure models merely create an echo that reinforces predispositions and leads to ignorant certainty. This is found in the social psychology of selective disclosure. Evidence considered in the literature as well as in [15] supports echo over bandwidth. Bandwidth and echo models represent a fundamental choice for theoretical models of trust. The lying in our model can be interpreted as selective disclosure in this context. It can be seen that passing on reputation as second hand information more likely creates an echo scenario. Passing on only direct experience more likely creates a bandwidth scenario.

In general, work in social sciences is more about data collection and interpretation than about modelling, although there are some. [10], for example, model interpersonal relationships using algebraic semigroups.

Another social phenomenon is that of herd behaviour. It has been considered in scenarios as diverse as voters behaviour at elections, fertility choices, technology decisions or hot topics in research. An economic approach is provided in [8]. The author analyzes a game theoretic model in which each decision maker looks at decisions made by previous decision makers in taking her own decision. The aim of the game for each player is to find the correct option. The fact that other players' influence is very strong, somewhat limits the applicability of the model.

6. CONCLUSIONS AND FUNDAMENTAL DIRECTIONS FOR RESEARCH

In this paper, we have motivated the importance of reputation for self-organized communication systems and provided a survey of distributed reputation system research in the various communities within computer science. We have also pointed at reputation research in the many disciplines outside computer science, focusing in particular on the social sciences. Moreover, we have presented results from our mathematical model addressing fundamental questions about the liars' impact on reputation.

It has become apparent that both definition and representation of reputation vary widely even within computer science. While it is debatable whether or not the same definition and representation should be used for all applications, a more coherent terminology would certainly be desirable. Moreover, it would be useful to have a coherent classification of reputation systems. Criteria (A1–8), (B1–3) and (C1–4) from the artificial intelligence, Internet-based P2P and Mobile Ad-Hoc Networks communities respectively (cf. Sections 3.1, 3.2 and 3.3) are relevant in all communities and could easily be combined into a more complete classification.

Based on a more coherent terminology, it would also be desirable to bring together the different strands of research, within computer science, but also between disciplines. This would avoid lots of reinventing the wheel that can be observed at the moment. The reputations research network⁴ went only some way towards this. In this paper, we have attempted to at least provide pointers to work on reputation in the different computer science communities as well as in social network analysis. A number of interesting examples of the successful combination of different disciplines (physics, economics, social sciences) can be found in [6].

Clearly, specific applications for distributed reputation systems are of crucial importance and many papers address various scenarios. However, it is also important not to get lost in the details. There are fundamental questions that are important in all these scenarios that should be addressed on a suitable level of abstraction. For example, the difference between passing on *Reputation* as opposed to *Direct Observations* falls into this category.

Finally, computer science research is often based on simulations, measurements or implementation and testing of prototypes. For example, prototype protocols are typically evaluated using a network simulator such as ns-2⁵ or GloMoSim⁶. Apart from game and graph theoretic investigations (cf. Section 3), there are comparatively few analytical studies although they often provide insight that is hard to obtain otherwise. An example of this is the stochastic process formulation of the model discussed in Section 4. Even though results are typically proven to be valid under clearly defined assumptions, some of which might be unrealistic, it is often the case that results are valid at least qualitatively even if the assumptions are violated. In the social sciences context in particular, such approaches are rare although it would be desirable to enhance predictive capabilities of social networks. Moreover, there might yet be other approaches (than game theoretic, graph theoretic and stochastic models) that have not been considered so far.

7. ACKNOWLEDGEMENTS

The authors would like to thank the organizers for their kind invitation. Thanks also to Sonja Buchegger, Radu Jurca and Wojciech Galuba for helpful discussions.

8. REFERENCES

- [1] K. Aberer and Z. Despotovic. Managing trust in a peer-2-peer information system. In *Proceedings of the Ninth International Conference on Information and Knowledge Management (CIKM)*, 2001.
- [2] K. Aberer, Z. Despotovic, W. Galuba, and W. Kellerer. *Computational Intelligence, Theory and Application*, chapter The complex facets of reputation and trust. Springer, to appear 2006.
- [3] E. Adar and B. Huberman. Free riding on gnutella. *First Monday*, 5(10), 2000.

⁴<http://databases.si.umich.edu/reputations/index.html>

⁵<http://www.isi.edu/nsnam/ns/>

⁶<http://pcl.cs.ucla.edu/projects/glomosim/>

- [4] J. Andreoni and J. H. Miller. Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica*, 70:737–753, 2002.
- [5] P. Antoniadis, C. Courcoubetis, and R. Mason. Comparing economic incentives in peer-to-peer networks. *Computer Networks*, 46(1):133–146, 2004.
- [6] P. Ball. *Critical mass: how one thing leads to another*. Farrar, Straus and Giroux, 2004.
- [7] G. Bamberg and K. Spremann. *Agency Theory, Information, and Incentives*. Springer, 1989.
- [8] A. V. Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, 170(3):797–817, 1992.
- [9] S. Bansal and M. Baker. Observation-based cooperation enforcement in ad hoc networks. Technical report, Arxiv preprint cs.NI/0307012, 2003.
- [10] G. R. Barnes, P. B. Cerrito, and I. Levi. A mathematical model for interpersonal relationships in social networks. *Social Networks*, 20:179–196, 1998.
- [11] S. Buchegger and J.-Y. Le Boudec. Performance Analysis of the CONFIDANT Protocol: Cooperation Of Nodes – Fairness In Dynamic Ad-hoc NeTworks. In *Proceedings of IEEE/ACM Symposium on Mobile Ad Hoc Networking and Computing (MobiHOC)*. IEEE, 2002.
- [12] S. Buchegger and J.-Y. Le Boudec. The effect of rumor spreading in reputation systems for mobile ad-hoc networks. In *Proceedings of WiOpt 2003: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, 2003.
- [13] S. Buchegger and J.-Y. Le Boudec. Self-policing mobile ad hoc networks by reputation systems. *IEEE Communications Magazine*, 43(7):101–107, July 2005.
- [14] R. S. Burt. The social capital of opinion leaders. *Annals of the American Academy of Political and Social Science*, 1999.
- [15] R. S. Burt. *Bandwidth and Echo: Trust, Information, and Gossip in Social Networks*, chapter in *Networks and Markets: Contributions from Economics and Sociology* (Casella, A. and J. E. Rauch (Edts.)). Russell Sage Foundation, 2001.
- [16] L. Buttyán and J.-P. Hubaux. Stimulating cooperation in self-organizing mobile ad hoc networks. *ACM/Kluwer Mobile Networks and Applications*, 8(5):579–592, 2003.
- [17] J. Chu, K. Labonte, and B. N. Levine. Availability and locality measurements of peer-to-peer file systems. In *ITCom: Scalability and Traffic Control in IP Networks, Proceedings of SPIE*, volume 4868, 2002.
- [18] R. Cornes and T. Sandler. *The Theory of Externalities, Public Goods, and Club Goods*. Cambridge University Press, 2nd edition, 1996.
- [19] C. Courcoubetis and R. R. Weber. *Pricing Communication Networks: Economics, Technology and Modelling*. Wiley Europe, 2003.
- [20] C. Courcoubetis and R. R. Weber. Asymptotics for provisioning problems of peering wireless LANs with a large number of participants. In *Proceedings of WiOpt 2004: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, 2004.
- [21] J. Crowcroft, R. Gibbens, F. Kelly, and S. Östring. Modelling incentives for collaboration in mobile ad hoc networks. *Performance Evaluation*, 57:427–439, 2003.
- [22] C. Dellarocas. The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science*, 49(10):1407–1424, 2003.
- [23] J. R. Douceur. *Peer-to-Peer Systems: First International Workshop, IPTPS 2002*, volume 2429 / 2002 of *Lecture Notes in Computer Science*, chapter The Sybil Attack, pages 251–260. Springer Berlin / Heidelberg, 2002.
- [24] U. Ebert and O. von dem Hagen. Altruism, redistribution and social insurance. *Review of Economic Design*, 5:365–385, 2000.
- [25] E. J. Friedman and P. Resnick. The social cost of cheap pseudonyms. *Journal of Economics and Management Strategy*, 10(2):173–199, 2001.
- [26] P. Golle, K. Leyton-Brown, I. Mironov, and M. Lillibridge. Incentives for sharing in peer-to-peer networks. *Lecture Notes in Computer Science*, 2232:75–87, 2001.
- [27] G. Hardin. The tragedy of the commons. *Science*, 162:1243–1248, 1968.
- [28] Q. He, D. Wu, and P. Khosla. SORI: a secure and objective reputation-based incentive scheme for ad hoc networks. In *IEEE Wireless Communications and Networking Conference (WCNC 2004)*, March 2004.
- [29] Y. Hu, D. Johnson, and D. Maltz. The dynamic source routing protocol for mobile ad hoc networks (DSR). <http://www.ietf.org/internet-drafts/draft-ietf-manet-dsr-10.txt>, July 2004.
- [30] J.-P. Hubaux, J.-Y. Le Boudec, S. Giordano, M. Hamdi, L. Blazević, L. Buttyán, and M. Vojnović. Towards Mobile Ad-Hoc WANS: Terminodes. In *Proceedings of IEEE WCNC 2000*, 2000.
- [31] A. Jösang, R. Ismail, and C. Boyd. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 2006.
- [32] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina. The eigentrust algorithm for reputation management in p2p networks. In *Proceedings of the Twelfth International World Wide Web Conference 2003*, 2003.

- [33] T. Karagiannis, A. Broido, N. Brownlee, k. claffy, and M. Faloutsos. Is P2P dying or just hiding? In *Proceedings of IEEE Globecom 2004: Global Internet and Next Generation Networks*, November 2004.
- [34] J. Kim and P. J. Bentley. Towards an artificial immune system for network intrusion detection: An investigation of dynamic clonal selection. In *Congress on Evolutionary Computation (CEC-2002)*, pages 1015–1020, 2002.
- [35] D. B. Klein, editor. *Reputation: Studies in the Voluntary Elicitation of Good Conduct. Economics, Cognition, and Society*. University of Michigan Press, 1997.
- [36] G. Lai and O. Wong. The tie effect on information dissemination: the spread of a commercial rumor in hong kong. *Social Networks*, 24:40–75, 2002.
- [37] T. Liggett. *Interacting Particle Systems*. Springer, New York, 1985.
- [38] H. Masum and Y.-C. Zhang. Manifesto for the reputation society. *First Monday*, 9(7), 2004.
- [39] P. Michiardi and R. Molva. Core: A Collaborative REputation mechanism to enforce node cooperation in Mobile Ad Hoc Networks. In *Proceedings of the Sixth IFIP Communication and Multimedia Security Conference 2002*, 2002.
- [40] N. Miller, P. Resnick, and R. Zeckhauser. Eliciting Informative Feedback: The Peer-Prediction Method. *Management Science*, 51:1359–1373, 2005.
- [41] H. H. S. Moerbeek and A. Need. Enemies at work: can they hinder your career. *Social Networks*, 25:67–82, 2003.
- [42] J. Mundinger. *Analysis of P2P Systems in Communication Networks*. PhD thesis, Cambridge University, August 2005.
- [43] J. Mundinger, S. Buchegger, and J.-Y. Le Boudec. Distributed reputation systems for internet-based peer-to-peer systems and mobile ad-hoc networks. *ERCIM News*, 63:19–20, October 2005.
- [44] J. Mundinger and J.-Y. Le Boudec. Analysis of a reputation system for mobile ad-hoc networks with liars. In *Proceedings of WiOpt 2005: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, pages 41–46, 2005.
- [45] J. Mundinger and J.-Y. Le Boudec. Analysis of a robust reputation system for self-organised networks. *European Transactions on Telecommunications, Special Issue on Self-Organisation in Mobile Networking*, 16(5):375–384, October 2005.
- [46] J. Mundinger and J.-Y. Le Boudec. The impact of liars on reputation in social networks. In *Proceedings of Social Network Analysis: Advances and Empirical Applications Forum*, Oxford, UK, July 2005.
- [47] D. C. North. *Institutions, Institutional Change and Economic Performance*. Cambridge University Press, 1990.
- [48] C. H. Papadimitriou. Algorithms, games, and the internet. In *33rd Annual ACM Symposium on Theory of Computing (STOC 2001)*, 2001.
- [49] P. Resnick, R. Zeckhauser, E. Friedman, and K. Kuwabara. Reputation systems: Facilitating trust in internet interactions. *Communications of the ACM*, 43(12):45–48, 2000.
- [50] T. Roughgarden. *Selfish Routing and the Price of Anarchy*. MIT Press, 2005.
- [51] J. Sabater and C. Sierra. Reputation and social network analysis in multi-agent systems. In *First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'02)*, 2002.
- [52] J. Sabater and C. Sierra. Review on computation trust and reputation models. *Artificial Intelligence Review*, 24:33–60, 2005.
- [53] P. A. Samuelson. The pure theory of public expenditure. *Review of Economics and Statistics*, 36(4):387–389, 1954.
- [54] S. Sarafijanovic and J.-Y. Le Boudec. An artificial immune system for misbehavior detection in mobile ad-hoc networks with virtual thymus, clustering, danger signal and memory detectors. *International Journal of Unconventional Computing*, 1:221–254, 2005.
- [55] S. Saroiu, P. K. Gummadi, and S. D. Gribble. A measurement study of peer-to-peer file sharing systems. In *Proceedings of Multimedia Computing and Networking (MMC� 2002)*, 2002.
- [56] J. Scott. *Social Network Analysis: A Handbook*. Sage Publications, 2nd edition, 2000.
- [57] A. Somayaji and S. Forrest. Automated response using system-call delays. In *Proceedings of the 9th USENIX Security Symposium*. The USENIX Association, 2000.
- [58] The Economist. Special report eBay. The Economist, June 11th, 2005, 2005.
- [59] A. Whitby, A. Jösang, and J. Indulska. Filterin out unfair ratings in bayesian reputation systems. In *Third International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'04)*, 2004.
- [60] L. Xiong and L. Liu. PeerTrust: Supporting Reputation-Based Trust in Peer-to-Peer Communities. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 16(7):843–857, July 2004.
- [61] B. Ycart. *Modèles et algorithmes markoviens*, volume 39. Springer Verlag, 2002.
- [62] B. Yu and M. Singh. Detecting deception in reputation management. In *Second International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'03)*, 2003.